OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/943,857

DATE: 01/19/2002
TIME: 11:07:09

Input Set : A:\08919-066001.TXT
Output Set: N:\CRF3\01192002\I943857.raw

ENTERED

<110> APPLICANT: Shaw, Jei-Fu
Lee, Guan-Chiun
Tang, Shye-Jye
<120> TITLE OF INVENTION: RECOMBINANT CANDIDA RUGOSA LIPASES
<130> FILE REFERENCE: 08919-066001
<140> CURRENT APPLICATION NUMBER: 09/943,857
<141> CURRENT FILING DATE: 2001-08-31
<160> NUMBER OF SEQ ID NOS: 11
<170> SOFTWARE: FastSEQ for Windows Version 4.0
<210> SEQ ID NO: 1
<211> LENGTH: 1469
<212> TYPE: DNA
<213> ORGANISM: Candida rugosa
<400> SEQUENCE: 1

gcccacgcca cctcgccaac ggcgacacca tcaccggtct caacgccatt gtcaacgaaa 60
agtttctcgg cataccgttt gccgagccgc ccgtgggcac tccgcttcaa gctcaacggc 120
cagcagttta cctacggccc gtgcatgcag atgaacccta tgggctcgtt tcattggtgc 180
tccagtccaa gatcttccaa gtggtgcttc ccaacgacga ggactgtctc accaccaggg 240
ccagtgctgg tctcccggtg atgctctgga tctttggcgg tgggtttgag cttggcggct 300
ccagcctctt tccaggagac cagatggtgg ccaagagcgt gctcatgggt aaaccggtga 360
tccacgtgag catgaactac cgcgtggcgt catgggggtt cttggccggc cccgacatcc 420
agaacgaagg cagcgggaac gccggcttgc atgaccagcg cttggccatg cagtgggtgg 480
cggacaacat tgctgggttt ggcggcgacc cgagcaaggt gaccatatac ggcgaggcgg 540
gcagcatgtc gacgtttgtg caccttgtgt ggaacgacgg cgacaacacg tacaacggca 600
agccgttgtt ccgcgccgcc atcatgcagg gctgcatggt gccggacccg gtggacggca 660
cgtacggcac cgagatctac aaccaggtgg tggcgtctgc cgggtgtggc agtgccagcg 720
acaagctcgc gtgcttgcgc ggcctttctc aggacacgtt gtaccaggcc acgagcgaca 780
cgcccggcgt gttggcgtac ccgtcgttgc ggttgtatct cccgcggccc gacggcacct 840
tcatcaccga cgacatgtat gccttggtgc gggacggcaa gtacgcacac gtgccggtga 900
tcatcggcga ccagaacgac gagggcactt tgtttgggct cttgaacgtg accacagatg 960
ctcaggcacg ggcgtacttc aagcagttca tccacgccag cgatgcggag atcgacacgt 1020
tgatggcggc gtacaccagc gacatcaccc agggtccgtt cgacaccggc atcttcaatg 1080
ccatcacccc gcagttcaaa cggatcgcgt tgcttggcga ccttgcgttc acgcttgcgc 1140
gtcgctactt cctcaactac taccagggcg gcaccaagta ctcgttctca agcagcttgg 1200
gttgcccgtc ttgggcacct tccacggcaa cgacatcatc tggcaggact acttggtggg 1260
cagcggcagt gtgatctaca acaacgcgtt cattgcgttt gccaacgacc tcgacccgaa 1320
caaggcgggc ttgtggacca actggcccac gtacaccagc agcagggcaa caacttgatg 1380
cagatcaacg gcttggggtt gtacaccggc aaggacaact tccgcccgga tgcgtacagc 1440
gccctctttt ccaacccgcc ttctttgtg 1469

<210> SEQ ID NO: 2
<211> LENGTH: 547
<212> TYPE: PRT
<213> ORGANISM: Candida rugosa

RAW SEQUENCE LISTING DATE: 01/19/2002
PATENT APPLICATION: US/09/943,857 TIME: 11:07:09

Input Set : A:\08919-066001.TXT
Output Set: N:\CRF3\01192002\I943857.raw

```
57 <400> SEQUENCE: 2
58 Ser Met Asn Ser Arg Gly Pro Ala Gly Arg Leu Gly Ser Val Pro Thr
59  1               5                  10                  15
60 Ala Thr Leu Ala Asn Gly Asp Thr Ile Thr Gly Leu Asn Ala Ile Val
61             20                  25                  30
62 Asn Glu Lys Phe Leu Gly Ile Pro Phe Ala Glu Pro Pro Val Gly Ser
63         35                  40                  45
64 Leu Arg Phe Lys Pro Pro Val Pro Tyr Ser Ala Ser Leu Asn Gly Gln
65     50                  55                  60
66 Gln Phe Thr Ser Tyr Gly Pro Ser Cys Met Gln Met Asn Pro Met Gly
67 65                  70                  75                  80
68 Ser Phe Glu Asp Thr Leu Pro Lys Asn Ala Leu Asp Leu Val Leu Gln
69                 85                  90                  95
70 Ser Lys Ile Phe Gln Val Val Leu Pro Asn Asp Glu Asp Cys Leu Thr
71             100                 105                 110
72 Ile Asn Val Ile Arg Pro Pro Gly Thr Arg Ala Ser Ala Gly Leu Pro
73         115                 120                 125
74 Val Met Leu Trp Ile Phe Gly Gly Gly Phe Glu Leu Gly Gly Ser Ser
75     130                 135                 140
76 Leu Phe Pro Gly Asp Gln Met Val Ala Lys Ser Val Leu Met Gly Lys
77 145                 150                 155                 160
78 Pro Val Ile His Val Ser Met Asn Tyr Arg Val Ala Ser Trp Gly Phe
79                 165                 170                 175
80 Leu Ala Gly Pro Asp Ile Gln Asn Glu Gly Ser Gly Asn Ala Gly Leu
81             180                 185                 190
82 His Asp Gln Arg Leu Ala Met Gln Trp Val Ala Asp Asn Ile Ala Gly
83         195                 200                 205
84 Phe Gly Gly Asp Pro Ser Lys Val Thr Ile Tyr Gly Glu Ser Ala Gly
85     210                 215                 220
86 Ser Met Ser Thr Phe Val His Leu Val Trp Asn Asp Gly Asp Asn Thr
87 225                 230                 235                 240
88 Tyr Asn Gly Lys Pro Leu Phe Arg Ala Ala Ile Met Gln Ser Gly Cys
89                 245                 250                 255
90 Met Val Pro Ser Asp Pro Val Asp Gly Thr Tyr Gly Thr Glu Ile Tyr
91             260                 265                 270
92 Asn Gln Val Val Ala Ser Ala Gly Cys Gly Ser Ala Ser Asp Lys Leu
93         275                 280                 285
94 Ala Cys Leu Arg Gly Leu Ser Gln Asp Thr Leu Tyr Gln Ala Thr Ser
95     290                 295                 300
96 Asp Thr Pro Gly Val Leu Ala Tyr Pro Ser Leu Arg Leu Ser Tyr Leu
97 305                 310                 315                 320
98 Pro Arg Pro Asp Gly Thr Phe Ile Thr Asp Asp Met Tyr Ala Leu Val
99                 325                 330                 335
100 Arg Asp Gly Lys Tyr Ala His Val Pro Val Ile Ile Gly Asp Gln Asn
101             340                 345                 350
102 Asp Glu Gly Thr Leu Phe Gly Leu Ser Ser Leu Asn Val Thr Thr Asp
103         355                 360                 365
104 Ala Gln Ala Arg Ala Tyr Phe Lys Gln Ser Phe Ile His Ala Ser Asp
105     370                 375                 380
```

RAW SEQUENCE LISTING DATE: 01/19/2002
PATENT APPLICATION: US/09/943,857 TIME: 11:07:09

Input Set : A:\08919-066001.TXT
Output Set: N:\CRF3\01192002\I943857.raw

```
106 Ala Glu Ile Asp Thr Leu Met Ala Ala Tyr Thr Ser Asp Ile Thr Gln
107 385                 390                 395                 400
108 Gly Ser Pro Phe Asp Thr Gly Ile Phe Asn Ala Ile Thr Pro Gln Phe
109                 405                 410                 415
110 Lys Arg Ile Ser Ala Leu Leu Gly Asp Leu Ala Phe Thr Leu Ala Arg
111             420                 425                 430
112 Arg Tyr Phe Leu Asn Tyr Tyr Gln Gly Gly Thr Lys Tyr Ser Phe Leu
113         435                 440                 445
114 Ser Lys Gln Leu Ser Gly Leu Pro Val Leu Gly Thr Phe His Gly Asn
115     450                 455                 460
116 Asp Ile Ile Trp Gln Asp Tyr Leu Val Gly Ser Gly Ser Val Ile Tyr
117 465                 470                 475                 480
118 Asn Asn Ala Phe Ile Ala Phe Ala Asn Asp Leu Asp Pro Asn Lys Ala
119                 485                 490                 495
120 Gly Leu Trp Thr Asn Trp Pro Thr Tyr Thr Ser Ser Ser Gln Ser Gly
121             500                 505                 510
122 Asn Asn Leu Met Gln Ile Asn Gly Leu Gly Leu Tyr Thr Gly Lys Asp
123         515                 520                 525
124 Asn Phe Arg Pro Asp Ala Tyr Ser Ala Leu Phe Ser Asn Pro Pro Ser
125     530                 535                 540
126 Phe Phe Val
127 545
129 <210> SEQ ID NO: 3
130 <211> LENGTH: 1532
131 <212> TYPE: DNA
132 <213> ORGANISM: Candida rugosa
134 <400> SEQUENCE: 3
135 gcccaccgcc aagctcgcca acggcgacac catcaccggt ctcaacgcca tcatcaacga      60
136 ggcgttcctc ggcattccct ttgccgagcc gccggtgggc aacctccgct tcaaggaccc     120
137 tgtgccgtac tctggctcgc tcaacggcca gaagttactt acggcccgtg catgcagcag     180
138 aaccccgagg gcacgtttga agagaacctt ggcaagacgg cactcgactt ggtgatgcag     240
139 tccaaggtgt tccaggcggt gcttccccag agtgaggact gcctcaccat caacgtggtg     300
140 cggccgccgg gcaccaaggc gggcgccaac ctcccggtca tgctctggat ctttggcggt     360
141 gggtttgaga tcggcagccc caccatcttc cctcccgccc agatggtcac caagagtgtg     420
142 ctcatgggca agccatcatc cacgtggccg tcaactaccg tgttgcctcg tgggggttct     480
143 tggctggtga tgacatcaag gccgagggca gcgggaacgc cggcttgaag gaccagcgtt     540
144 tgggcatgca gtgggtggca gacaacattg ccgggttcgg cggcgacccg agcaaggtga     600
145 catctttggc gaggcgggca gcatgtccgt gttgtgccac ctcatctgga acgacggcga     660
146 caacacgtac aagggcaagc cgttgttccg cgcgggcatc atgcagggag ccatggtgcc     720
147 ggacccggtg gacggcacgt acggcaacga gatctacgac ctctttgtct cgagtgctgg     780
148 ctgtggcagc gccagcgaca agctcgcgtg cttgcgcagt gcgagcgaca ccttgctcga     840
149 tgccaccaac aacactcctg ggttcttggc gtactcctcg ttgcggttgt actcccggcc     900
150 cgacggcaag aacatcaccg atgacatgta caagttggtg cgcgacggca agtatgcaag     960
151 cgttcccgtg atcattggcg accagaacga cgagggcacc atctttggct cttgaacgtg    1020
152 accacgaatg ctcaggcccg tgcttacttc aagcagttca tccacgccag cgacgcggag    1080
153 atcgacacct tgatggcggc gtaccccccag gacatcaccc agggtccgtt cgacacggtt    1140
154 caacgctcac cccgcagttc aagagaatcg cggtgctcgg cgaccttgca ttcatccacg    1200
155 cccgccgcta cttcctcaac cacttccagg gcggcaccaa gtactcgttc ctcaagcagc    1260
156 tcgggttgcc aatcatgggc accttccatg ccaacgacat tgtgtggcag gactacttgt    1320
```

RAW SEQUENCE LISTING DATE: 01/19/2002
PATENT APPLICATION: US/09/943,857 TIME: 11:07:09

Input Set : A:\08919-066001.TXT
Output Set: N:\CRF3\01192002\I943857.raw

```
157 tgggaagcgg cagcgtcatc tacaacaacg cgtttatcgc gttcgccacc gacttggacc    1380
158 ccaacaccgc ggggttgttg gtgaactggc ccaagtacac cagcagccag ggcaacaact    1440
159 tgatgatgat caacgccttg ggcttgtaca ccggcaagga caacttccgc accgctggct    1500
160 acgacgcgtt gatgaccaac ccgttctttg tg                                  1532
162 <210> SEQ ID NO: 4
163 <211> LENGTH: 547
164 <212> TYPE: PRT
165 <213> ORGANISM: Candida rugosa
167 <400> SEQUENCE: 4
168 Ser Met Asn Ser Arg Gly Pro Ala Gly Arg Leu Gly Ser Val Pro Thr
169  1               5                  10                  15
170 Ala Lys Leu Ala Asn Gly Asp Thr Ile Thr Gly Leu Asn Ala Ile Ile
171             20                  25                  30
172 Asn Glu Ala Phe Leu Gly Ile Pro Phe Ala Glu Pro Pro Val Gly Asn
173         35                  40                  45
174 Leu Arg Phe Lys Asp Pro Val Pro Tyr Ser Gly Ser Leu Asn Gly Gln
175     50                  55                  60
176 Lys Phe Thr Ser Tyr Gly Pro Ser Cys Met Gln Gln Asn Pro Glu Gly
177 65                  70                  75                  80
178 Thr Phe Glu Glu Asn Leu Gly Lys Thr Ala Leu Asp Leu Val Met Gln
179                 85                  90                  95
180 Ser Lys Val Phe Gln Ala Val Leu Pro Gln Ser Glu Asp Cys Leu Thr
181             100                 105                 110
182 Ile Asn Val Val Arg Pro Pro Gly Thr Lys Ala Gly Ala Asn Leu Pro
183         115                 120                 125
184 Val Met Leu Trp Ile Phe Gly Gly Gly Phe Glu Ile Gly Ser Pro Thr
185     130                 135                 140
186 Ile Phe Pro Pro Ala Gln Met Val Thr Lys Ser Val Leu Met Gly Lys
187 145                 150                 155                 160
188 His Ile Ile His Val Ala Val Asn Tyr Arg Val Ala Ser Trp Gly Phe
189                 165                 170                 175
190 Leu Ala Gly Asp Asp Ile Lys Ala Glu Gly Ser Gly Asn Ala Gly Leu
191             180                 185                 190
192 Lys Asp Gln Arg Leu Gly Met Gln Trp Val Ala Asp Asn Ile Ala Gly
193         195                 200                 205
194 Phe Gly Gly Asp Pro Ser Lys Val Thr Ile Phe Gly Glu Ser Ala Gly
195     210                 215                 220
196 Ser Met Ser Val Leu Cys His Leu Ile Trp Asn Asp Gly Asp Asn Thr
197 225                 230                 235                 240
198 Tyr Lys Gly Lys Pro Leu Phe Arg Ala Gly Ile Met Gln Ser Gly Ala
199                 245                 250                 255
200 Met Val Pro Ser Asp Pro Val Asp Gly Thr Tyr Gly Asn Glu Ile Tyr
201             260                 265                 270
202 Asp Leu Phe Val Ser Ser Ala Gly Cys Gly Ser Ala Ser Asp Lys Leu
203         275                 280                 285
204 Ala Cys Leu Arg Ser Ala Ser Ser Asp Thr Leu Leu Asp Ala Thr Asn
205     290                 295                 300
206 Asn Thr Pro Gly Phe Leu Ala Tyr Ser Ser Leu Arg Leu Ser Tyr Leu
207 305                 310                 315                 320
```

RAW SEQUENCE LISTING DATE: 01/19/2002
PATENT APPLICATION: US/09/943,857 TIME: 11:07:09

Input Set : A:\08919-066001.TXT
Output Set: N:\CRF3\01192002\I943857.raw

```
208 Pro Arg Pro Asp Gly Lys Asn Ile Thr Asp Asp Met Tyr Lys Leu Val
209                 325                 330                 335
210 Arg Asp Gly Lys Tyr Ala Ser Val Pro Val Ile Ile Gly Asp Gln Asn
211             340                 345                 350
212 Asp Glu Gly Thr Ile Phe Gly Leu Ser Ser Leu Asn Val Thr Thr Asn
213         355                 360                 365
214 Ala Gln Ala Arg Ala Tyr Phe Lys Gln Ser Phe Ile His Ala Ser Asp
215     370                 375                 380
216 Ala Glu Ile Asp Thr Leu Met Ala Ala Tyr Pro Gln Asp Ile Thr Gln
217 385                 390                 395                 400
218 Gly Ser Pro Phe Asp Thr Gly Val Leu Asn Ala Leu Thr Pro Gln Phe
219                 405                 410                 415
220 Lys Arg Ile Ser Ala Val Leu Gly Asp Leu Ala Phe Ile His Ala Arg
221             420                 425                 430
222 Arg Tyr Phe Leu Asn His Phe Gln Gly Gly Thr Lys Tyr Ser Phe Leu
223         435                 440                 445
224 Ser Lys Gln Leu Ser Gly Leu Pro Ile Met Gly Thr Phe His Ala Asn
225     450                 455                 460
226 Asp Ile Val Trp Gln Asp Tyr Leu Leu Gly Ser Gly Ser Val Ile Tyr
227 465                 470                 475                 480
228 Asn Asn Ala Phe Ile Ala Phe Ala Thr Asp Leu Asp Pro Asn Thr Ala
229                 485                 490                 495
230 Gly Leu Leu Val Asn Trp Pro Lys Tyr Thr Ser Ser Ser Gln Ser Gly
231             500                 505                 510
232 Asn Asn Leu Met Met Ile Asn Ala Leu Gly Leu Tyr Thr Gly Lys Asp
233         515                 520                 525
234 Asn Phe Arg Thr Ala Gly Tyr Asp Ala Leu Met Thr Asn Pro Ser Ser
235     530                 535                 540
236 Phe Phe Val
237 545
239 <210> SEQ ID NO: 5
240 <211> LENGTH: 1548
241 <212> TYPE: DNA
242 <213> ORGANISM: Candida rugosa
244 <400> SEQUENCE: 5
245 gcccactgcc acgctcgcca acggcgacac catcaccggt ctcaacgcca tcatcaacga      60
246 ggcgttcctc ggtattccct ttgctcagcc gccggtgggc aacctccgct tcaagccgcc     120
247 tgtgccgtac tcggcgtctc tcaatggtca gaagtttact gtatggccct gtgcatgcag     180
248 atgaacccat tgggcaactg ggactcctcg cttcccaagg ctgccatcaa cttgatgcag     240
249 tccaagctct tccaggcggt gcttcctaac ggcgaggact gtctcaccat caacgtggtg     300
250 cggccgggca ccaagccggg tgccaacctc cccgtgatgg tgtggatttt tggcggcggg     360
251 tttgaggttg gcggctccag tctcttccct cccgcacaga tgatcaccgc cagcgtgctt     420
252 atgggcaagc ccatcatcca cgtgagcatg aactaccgcg ttgcttcgtg ggggttcttg     480
253 gctggtccag acatcaaggc cgagggcagc gggaacgccg gtttgcacga ccaacgcttg     540
254 ggtttgcagt gggtggcgga caacattgcc gggttcggcg gcgacccgtc caaggtgacc     600
255 atctttggtg agggcgggca gcatgtcggt aatgtgtcag ctcctctgga acgacggcga     660
256 caacacgtac aacggcaagc cgttgttccg tgccgccatc atgcaggggg ccatggtgcc     720
257 gggacccggt ggatgggccc tacggcacgc agatctacga ccaggtggtt gcttcagccg     780
258 gctgtggcag tgccagcgac aagctcgcgt gcttgcgcag catcgaacga caaactcttc     840
```

VERIFICATION SUMMARY DATE: 01/19/2002
PATENT APPLICATION: US/09/943,857 TIME: 11:07:10

Input Set : A:\08919-066001.TXT
Output Set: N:\CRF3\01192002\I943857.raw